Quantum communication and quantum networks

Jens Eisert
Dahlem Center for Complex Quantum Systems
Freie Universität Berlin

Contents

1 What this lecture is about

1.1 A short history of cryptography

Ideas of cryptography and secret communication are presumably about as old as mankind is. There are many reasons why one would like to communicate with a legitimate recipient while making sure at the same time that nobody else listens in to the conversation. This feature of communication is intricately intertwined with rather obvious features of human behaviour. For this reason it may not be a huge surprise that the history of cryptography reads like a crime story in the first place.

Examples of applications of cryptography from the more recent past (viewed from the perspective of the history of mankind, that is) include the cryptographic encoding of messages by a scytale, a device used as a cipher by the ancient Greeks and Spartans during military campaigns, first mentioned by the Greek poet Archilochus, who lived in the 7th century BC. It already features many aspects of a modern cryptographic scheme. It consists of a cylinder with a strip of parchment wound around it on which a message is written. The encryption arises from the fact that both the sender and the legitimate receiver share the cylinder. Once this is available, one can wind the parchment around it to generate a perfectly readable message. Without it, the message seems scrambled. The key point is that while two legitimate parties share the same object (a cylinder in this case, so a key in more modern terms), illegimate users would not have access to this object. While this idea gives rise to a code that can obviously be broken, it has a security level that is presumably sufficient to reflect combat situations in the ancient world.

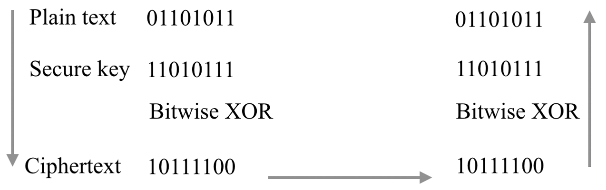
Turning to more recent events, it is well known that the fates of history in times of the second world war have been deeply intertwined with the history of secure communication. For example, Admiral Isoroku Yamamoto, the leading military commander of the Japanese Navy during World War II and the architect for the attack on Pearl Harbor, announced his advent to the front line base on the island of Bougainville to boost morale – of course strictly encrypted, that is. Only that it was not sufficiently encrypted after all. The encryption system used – the Japanese Naval Cipher JN-25D in this case – was intercepted and with some effort successfully decrypted by US naval intelligence units. By the time, Yamamoto was arriving, the US was already there.

Maybe even more prominently, during World War II, efforts of encryption and efforts of deciphering messages had a decisive impact. Submarines obviously make sense only if their precise location can be concealed. The *Enigma machine* was the machine in the focus of a number of pivotal events. It was an electro-mechanical rotor cipher machine, invented by the German engineer Arthur Scherbius at the end of World War I and later developed into various variants, that were developed in the early 20th century to protect commercial, diplomatic and military communication. What the Enigma does, basically, is to transform each letter into a product of permutations. Unlike the previously mentioned cryptosystems, it required serious effort to break the code. An early version of the Enigma was broken by the Polish General Staff's Cipher Bureau in December 1932. Later versions used by Nazi Germany could initially not be deciphered; early on in WW II, the British Government Code and Cypher School at Bletchley Park built up an extensive cryptanalytic capability to break later versions of the machine. Alan Turing, a Cambridge University mathematician and logician and the inventor of the famous paradigmatic Turing machine, provided much of the key insights that eventually led to the breaking the naval Enigma, which had a major influence on the naval war. Once messages sent by submarines could be deciphered, the advantage of submarines was gone, with significant implications on how the war developed. That is to say, Alan Turing and his team at Bletchley Park had a major contribution to allied victory (a state of affairs that was later less appreciated when he was very badly treated, but this is a different matter, intertwined with another historical development).

These examples are mentioned only to highlight how the history of cryptography – as a history of code making and code breaking – is intimately intertwined with important events in history. By no means is this supposed to mean that the use of cryptography is confined to the military realm. Quite to the contrary, the use of cryptography is permeant to many aspects of our modern lives, in fact, it is ubiquitous. Whenever one uses WhatsApp, `https`, or any instance of internet banking, one resorts to a cryptographic scheme. Secure communication has become a pillar of how we communicate. A truly fascinating story of the history of cryptography and how it is intertwined with human behaviour that is a great read is presented in Ref. [58].

1.2 One-time pads

Getting more concrete: One can communicate securely if two parties share the same key. In retrospect this may seem obvious, at the same time it used to be far from clear. The one-time pad was developed by Gilbert Vernam in 1917 [61], proving that there is an absolutely secure coding scheme which is secure against eavesdroppers with unlimited computational power. In the one-time pad, a plaintext is encoded making use of a secret key (a pad, for that matter) that has the same length as the plaintext itself. The very same key is also employed by the legitimate receiver to decode the message. Given that one makes use of the key only once, the encryption scheme is absolutely secure, a statement that has later been proven by Claude Shannon [54]. In modern cryptographic systems (such as the *Data Encryption Standard (DES)* [8] and the *Advanced Encryption Standard (AES)* [15]) now used widely, shorter keys are being made use of to encrypt longer messages, for obvious pragmatic reasons. Such an approach uses fewer resources, but at the same time is not to the same extent provably secure as the one-time pad is. In any case, ultimately, at the heart of the matter is how to establish a secure key in the first place.



1.3 Public key distribution schemes

The most commonly used scheme is based on so-called *public key cryptographic protocols*, prominently the famous *RSA scheme* named after Ron Rivest, Adi Shamir, and Leonard Adleman. This ingenious idea has actually been invented twice, once by RSA and once by James H. Ellis. Ellis was a British engineer and cryptographer who in 1970 also invented a public key distribution scheme while working at the Government Communications Headquarters (GCHQ) in Cheltenham. At the time his results were kept secret; they became available only later after the embargo had been lifted. In public key distribution schemes, a message receiver, now and later on referred to as Bob, prepares two different cryptographic keys. One that is public and one that is private. Subsequently, Bob broadcasts the public key through an authenticated channel so that everyone who listens to this channel can acquire a copy of the public key. There is no requirement whatsoever to keep this public key secret. The original sender of the message, referred to as Alice, encodes her message with the public key from Bob and then sends out the encrypted message through a public insecure channel. The algorithm is set up in such a fashion so that the message encrypted with the public key can only be decrypted in conjunction with the private key.

Public key systems are widely used, basically any cryptographic scheme one encounters in electronic communication is based on a public key cryptographic scheme. The RSA scheme is practically secure, with a security level depending on the key length. Unfortunately, its security has not been proven. It rests on the existence of *one-way functions*: The multiplication is in P, while factoring is contained in NP.

The *RSA algorithm* involves basically four steps: key generation, key distribution, encryption and decryption. The core idea is the observation that it is practically possible to identify three very large positive integers e, d and n with the property that the modular exponentiation for all integers m with $0 \leq m < n$ satisfies

$$(m^e)^d = m(\operatorname{mod} n) \tag{1}$$

and that even knowing e and n or even m it can be extremely difficult to find d. RSA involves a public key and a private key. e basically takes the role of the public key, d is kept as the private key exponent. Primality test [1], the decision problem that asks whether a given number is a prime number or not, used to be in NP, until a probabilistic algorithm in BPP became known, and later the algorithm was de-randomized to an algorithm in P (look for the Miller-Rabin primality test and Solovay-Strassen primality test). A proof of $\mathsf{P} = \mathsf{NP}$ would indeed prove that one-way functions do not exist, shaking the basis on which RSA rests. This would imply that there cannot be proven security in public key distribution schemes. However, the precise practical implications would depend on the specifics of the argument. For example, if the proof of $\mathsf{P} = \mathsf{NP}$ was not constructive, then this proof would not give advice on how to actually break the key.

1.4 Quantum computers potentially breaking public key schemes

In any case, there is no denying that the lack of provable security poses a significant security risk. RSA itself was a highly unexpected discovery, and one should hence not rule out the possibility that someone could find an efficient factoring algorithm and thus compromise most public cryptographic systems. What is more, a *quantum computer* [44] can solve factoring in polynomial time (*Shor's algorithm* [55] provides a quantum algorithm for factoring the runtime of which scales polynomially in the length of the input - it is in BQP in the language of computational complexity). Large-scale quantum computers do not exist yet, but the development is fast. In 2016, IBM made a 16 qubit cloud quantum computer publicly available as a cloud service based on superconducting circuits, which has been characterized using randomized benchmarking and developed into a 50 qubit machine in 2018. More recently still, Google announced the 128 qubit Brizzlecone chip, based on a similar architecture. These devices are still way too small (and too noisy) to pose a security risk. But their development is fast and the case for quantum computing is open. And indeed, large-scale quantum computers could break essentially all RSA based cryptographic schemes used today over night.

1.5 What quantum key distribution can deliver

Quantum key distribution is different. Its security on the level of the scheme is mathematically proven. Its security is based on very fundamental physical laws of nature. These are the laws of quantum mechanics. *Quantum mechanics* is the theory of the world at the small scale: That of atoms, ions and light quanta. But since the macroscopic world is ultimately built from such building blocks, it equally applies to the macroscopic world: It is the best physical theory of nature that we have today. In quantum key distribution, one envisions to make use of constituents in which the quantum features are most manifest. Practically speaking, one sends single photons (excitations of light modes), weak pulses or Gaussian light through fibres (the same kind of fibres that are used by the Telekom) or free space, even via satellites.

Ultimately, the security is rooted in structure elements of quantum mechanics: One cannot learn about the unknown quantum state of a quantum systems without disturbing the state. There are trade-offs: One can perform a gentle measurement, learn very little and at the same time disturb very little. And one can do hard projective measurements. But there is no way one can obtain some information about an unknown quantum state without changing the same state to some extent. An implication of this feature is that quantum information cannot be *copied* or *cloned*, as one commonly says in this context. It is impossible to build a machine that takes a physical system in an unknown quantum state and produces two quantum systems in the very same state. If one could do disturbance-free measurements, that would be possible, but the *no cloning* feature of quantum mechanics [18, 67] forbids that. We will see that this is a simple consequence of the linearity of quantum mechanical laws. Quantum key distribution is no far-fetched dream: It is already reality. One can commercially buy quantum cryptographic devices: The company IDQuantique is only one out of many offering such products. It has been one of the early successes of the field of quantum cryptography to implement a BB84 scheme (the simplest and most used scheme for quantum key distribution that we will discuss soon) making use of an installed optical fibre cable linking Geneva and Nyon over 23 km through Lake Geneva in 1995, at remarkably low quantum bit error rates [42, 43]. This effort basically started the development of long-distance quantum key distribution. In the meantime, satellite-based quantum key distribution is being pursued.

Why is not all modern cryptography done via quantum key distribution and it is still a market niche? This has various reasons. The core reason is that reliable quantum key distribution over arbitrary distances is still hindered by serious technological obstacles. One needs to build so-called *quantum repeaters* to compensate for losses, in order to maintain security in the presence of realisticly high noise levels. There is significant progress in this direction, but fully fledged quantum repeaters have not been implemented yet. This means that quantum key distribution is still confined to relatively short distances. Then, it is a marketing issue: The market may well grow a lot if people realize that the security claim in quantum key distribution is very different from that in public key distribution. Such processes take time. The BMBF (Bundesministerium für Bildung und Forschung) has "bug-proof communication" on its web page as one of the strategic aims, and indeed, there is a large scale project on realizing quantum repeaters by the BMBF, called Q.Link-X (which we are part of). It should be clear that quantum key distribution is no science fiction, but an important technology of tomorrow.

1.6 Some further reading

Quantum cryptography is a young field, but not that young, and the literature on the subject is extensive. It is a sub-field of *quantum information science*, the field of research exploring applications of single quantum systems to address tasks of information processing. The following list provides some hints at good literature in the field, even though this list makes no claim of completeness in any sense whatsoever.

- This much seen and cited review article on "quantum cryptography" [26] dates back to 2002 and is hence no longer entirely new. What is more, it puts a strong emphasis on practical implementations and not so much on mathematical details. However, it remains an excellent source for information on how one can realistically implement quantum key distribution schemes. It addresses physicists much more than mathematicians.

- The review on "cryptographic security of quantum key distribution" [48] takes a very different perspective. Here, the mathematical foundations of security proofs are at the heart of the matter.
- A brief but excellent overview for the impatient, "a brief introduction of quantum cryptography for engineers" can be found here [49].
- The topic of this course is basically quantum information theory. The text book on the subject matter, on "quantum computation and quantum information" [44], dates back to the year 2000, but is still surprisingly fresh. The topics on the stabilizer formalism that we will cover here are nicely explained there.
- Having said that, issues of quantum communication are studied comparably little in that text book. This omission is fixed by the text book [66], actually published by the same publisher, Cambridge University Press. It beautifully explains all notions of channel capacities that we will discuss here.
- A significant proportion of the course will be dedicated to multipartite quantum networks beyond point-to-point architectures. Here, interesting *graph problems* come into play, related to routing and scheduling as well as questions of computational complexity. Here is a strong link to the other two courses of this thematic semester of MATH+. This is a young field that is just emerging. Literature on this aspect can be found here [16, 17, 20, 22, 23, 28], with background in the relevant graph theory being presented here [9, 14].

2 Elements of quantum mechanics and a bit more

Quantum mechanics is a physical theory. Obviously, there are entire courses with more than double the number of lectures dedicated to elementary quantum mechanics. Last term, I was teaching another four hour lecture on "advanced quantum mechanics", and again we have only been looking at the tip of the iceberg. That is to say, we will have to keep the background as minimal as possible, to get going with our main theme.

2.1 Quantum bits

Classical bits can take the values 0 and 1 only. This is the commonly used basic unit of information, reflecting an on and off state of a basic cell. The state space of a classical bit is the straight line segment, reflecting a "mixture" or a convex combination 0 and 1. If the probability of having 0 is p_0 and that of having one 1 is p_1, then the state of the system is given by a vector $(p_0, p_1) \in \mathbb{R}^2$ with

$$p_0, p_1 \geq 0 \tag{2}$$

normalized as

$$p_0 + p_1 = 1. \tag{3}$$

This may be a bit of an overloaded way of putting it: But this is a convex set, a simplex in fact, and $(1, 0)$ and $(0, 1)$ are the extreme points of this set. Probabilistic mixtures take values in the interior of the set.

For quantum systems, we basically need to know that the state space is much bigger. This is the heart of the matter why quantum systems are more powerful than classical systems when it comes to applications in information processing. The equivalent of the bit is the *quantum bit*, in short *qubit*. Its state space is no longer a straight line segment, but can be represented as a ball, the *Bloch ball*. It generalizes probability distributions to matrices

$$\rho = \left[\begin{array}{cc} p_0 & c \\ c^* & p_1 \end{array} \right] \in \mathbb{C}^{2\times 2}. \tag{4}$$

Eq. (2) is being replaced by the constraint that ρ is positive semi-definite,

$$\rho \geq 0, \tag{5}$$

that it is normalized as

$$\mathrm{tr}(\rho) = 1. \tag{6}$$

Such a matrix ρ is called *density matrix* or simply the *quantum state* of the qubit. Since this density matrix is obviously Hermitian, its main diagonal elements are clearly real, and they are positive by virtue of Eq. (5). In fact, due to Eq. (6), they can be identified with a classical probability distribution (p_0, p_1). In fact, diagonal density operators can be identified with finite probability distributions.

But there is more to a quantum state: There is now an off-diagonal element $c \in \mathbb{C}$ of ρ. This may be innocent looking, but makes a big difference. One can no longer interpret a quantum state as a classical alternative. It is not in a probabilistic mixture of 0 or 1. In fact, the off-diagonal blocks signify a superposition, the qubit can be in "0 and 1 at the same time". It is common for quantum systems to be in such superpositions, even if our everyday intuition may find this alien or strange.

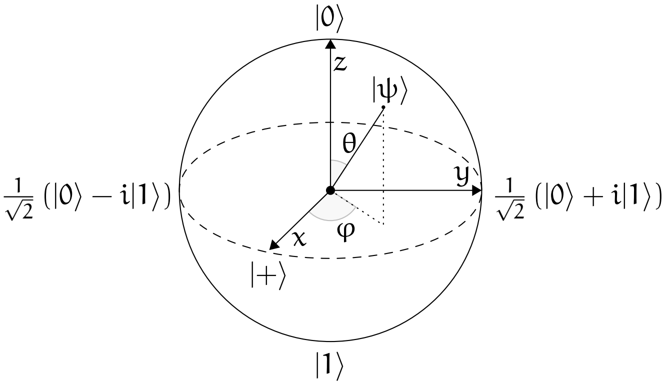
This becomes even more manifest when looking at the extreme points of the state space of a qubit. It is no longer a simplex. Extreme points can be written as complex vectors *("state vectors")*

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{7}$$

with $\alpha, \beta \in \mathbb{C}$ satisfying $|\alpha|^2 + |\beta|^2 = 1$. $|0\rangle$ and $|1\rangle$ are basic vectors of $\mathbb{C}^2$, written as "kets". These vectors form the basis of a vector space $\mathbb{C}^2$. That is to say, the pure states, the extreme points, of qubits are elements of a vector space, and any linear superposition gives rise to a legitimate pure state. The corresponding dual vectors are commonly written as $\langle 0|$ and $\langle 1|$, and referred to as "bras". Standard scalar product hence become "bra-kets", so brackets. The respective rank-1 projections onto $|0\rangle$ and $|1\rangle$ are then given by $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$. These basis vectors are isomorphic to density operators as

$$|0\rangle \approx |0\rangle\langle 0| \approx \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right], \; |1\rangle \approx |1\rangle\langle 1| \approx \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \right]. \tag{8}$$

It is both common to use the vector notation for pure states (i.e., extreme points of the set) as well as density matrices. It will depend on the context what is more natural to use. So far, we have already learned that a qubit has a larger state space than a simple bit, reflecting the superposition principle that is not present classically in the same fashion. For those students with a quantum mechanics background, this looks all very basic, for those with no quantum background this may require some digestion.



$|0\rangle$
z
$|\psi\rangle$
θ
$\frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$
y
$\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$
x
φ
$|+\rangle$
$|1\rangle$

Figure 1: The state space of qubits can be represented by the so-called Bloch sphere: Eq. 7 can be rewritten as $|\psi\rangle = e^{i\varphi_0} \cos \frac{\theta}{2}|0\rangle + e^{i\varphi_1} \sin \frac{\theta}{2}|1\rangle$. Since global phases do not have any observable effect in quantum mechanics we write w.l.o.g. $|\psi\rangle = \cos \frac{\theta}{2}|0\rangle + e^{i\varphi} \sin \frac{\theta}{2}|1\rangle$ with $\varphi = \varphi_1 - \varphi_0$.

2.2 Quantum state vectors of composite quantum systems

So far, we have learned what the state space of a single qubit is. We will not delve into general state spaces in quantum mechanics. For our purposes we only need to know what happens if we have many qubits at hand. The mathematical structure reflecting a composition of quantum systems is that of a *tensor product.*[1] The pure states, the expreme points of n qubits constitute the vector space $\mathcal{H} := \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 = (\mathbb{C}^2)^{\otimes n}$, reflecting the fact that an arbitrary state vector can be written as

$$|\psi\rangle = \alpha_{0,\ldots,0,0}|0\rangle \otimes \cdots \otimes |0\rangle \otimes |0\rangle + \alpha_{0,\ldots,0,1}|0\rangle \otimes \cdots \otimes |0\rangle \otimes |1\rangle + \ldots \alpha_{1,\ldots,1,1}|1\rangle \otimes \cdots \otimes |1\rangle \otimes |1\rangle. \quad (12)$$

The new basis of $\mathcal{H}$ is hence

$$\mathcal{B} := \{|i_1\rangle \otimes \cdots \otimes |i_{n-1}\rangle \otimes |i_n\rangle,\ i_1, \ldots, i_n \in \{0,1\}\}. \quad (13)$$

Since the many tensor products can be clumsy, one often writes $|0, \ldots, 0, 0\rangle$ instead of $|0\rangle \otimes \cdots \otimes |0\rangle \otimes |0\rangle$. Again, an *arbitrary superposition* of basis vectors as in Eq. (12) is a legitimate state vector corresponding to a pure state. This reflects the situation that a collection of qubits can – in a sense – be in “all classical alternatives at once”. This idea is also at the heart of quantum computing, in that a register is simultaneously manipulated in a superposition state reflecting several inputs “at once”. The precise functioning is subtle and more complicated than that, but this statement already creates the right mental image to see what this is about. We come back to this at the end of this section.

[1] Basic linear algebraic properties of the tensor product are taken for granted in this course. E.g., tensor products satisfy

$$\begin{aligned} |\psi\rangle \otimes |\omega\rangle + |\phi\rangle \otimes |\omega\rangle &= (|\psi\rangle + |\phi\rangle) \otimes |\omega\rangle, && (9)\\ |\omega\rangle \otimes |\psi\rangle + |\omega\rangle \otimes |\phi\rangle &= |\omega\rangle \otimes (|\psi\rangle + |\phi\rangle), && (10)\\ \alpha|\psi\rangle \otimes |\phi\rangle &= (\alpha|\psi\rangle) \otimes |\phi\rangle = |\psi\rangle \otimes (\alpha|\phi\rangle) && (11) \end{aligned}$$

for $\alpha \in \mathbb{C}$ and $|\psi\rangle$, $\phi\rangle$, $|\omega\rangle$ being state vector of their respective vector spaces.

2.3 Quantum state spaces as convex sets of positive semi-definite operators

In the same way, general *density operators* or *quantum states* are positive semi-definite matrices over this vectors space of dimension $\dim(\mathcal{H}) = 2^n =: d$.

Definition 1 (Quantum states of n qubits) *A general quantum state of a system of n qubits is given by a bounded operator ρ over the vector space $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ that is positive semi-definite and normalized as*

$$\rho \geq 0, \quad \operatorname{tr}(\rho) = 1. \tag{14}$$

The convex set of such operators is referred to as the state space $\mathcal{S}(\mathcal{H}) \subset \mathbb{C}^{d\times d}$. *The extreme points satisfy* $\operatorname{tr}(\rho^2) = 1$ *and correspond to vectors reflecting "pure states", they can be written as vectors $|\psi\rangle \in \mathcal{H}$ in that vector space, normalized as $\langle\psi|\psi\rangle = 1$.*

The set $\mathcal{S}(\mathcal{H})$ is indeed a *convex set*: If $\rho_1 \in \mathcal{S}(\mathcal{H})$ and $\rho_2 \in \mathcal{S}(\mathcal{H})$, then the straight line segment

$$\lambda\rho_1 + (1-\lambda)\rho_2 \in \mathcal{S}(\mathcal{H}) \tag{15}$$

is again contained in state space. That is to say, the *mixing* (i.e., the convex combination) of two quantum states ρ_1 and ρ_2 gives again rise to a valid quantum state. The interpretation of mixing is basically the same as that of convex combinations of probability distributions.

2.4 Projective measurements

We will turn to general measurements in the next section. For now, we will discuss projective selective so-called von-Neumann measurements only. Measurement in quantum mechanics are intrinsically *random*. This is another key feature of quantum mechanics. Upon performing a measurement, the theory only gives an indication on the probability of getting a certain outcome. But it is utterly silent about the specific outcome that is being obtained. This may be a bit of an odd feature of quantum mechanics, but it runs deep. In a way, one can say that the randomness of quantum mechanics is absolute, a statement that can be made precise by resorting to *Bell's theorem*. Any measurement prescription in quantum mechanics should presumably say what *property* is measured, but surely has to assign the *probability* of getting the respective outcome. Let us assume that before the measurement, the quantum state is given by $\rho \in \mathcal{S}(\mathcal{H})$, and call $d = \dim(\mathcal{H})$.

- In the simplest possible measurement, a von-Neumann measurement, one measures as a property a so-called *observable*, a Hermitian operator $A = A^\dagger$.
- The outcomes of the measurement are associated with a basis $\mathcal{B} = \{|\psi_j\rangle, j = 0, \dots, d-1\}$ of eigenvectors in $\mathcal{H}$.
- The probability of getting the outcome labeled $j = 0, \dots, d-1$ is given by $\mathbb{P}(j) = \langle\psi_j|\rho|\psi_j\rangle$.

Let us make this a definition.

Definition 2 (Von-Neumann measurements) *A von-Neumann measurement of an observable $A = A^\dagger$ featuring a non-degenerate spectrum of a quantum system equipped with a vector space $\mathcal{H}$ receives the outcomes labeled $j = 0, \dots, d-1$ with probability*

$$\mathbb{P}(j) = \langle \psi_j | \rho | \psi_j \rangle, \tag{16}$$

where $\mathcal{B} = \{|\psi_j\rangle, j = 0, \dots, d-1\}$ is an orthonormal basis of eigenvectors in $\mathcal{H}$. Immediately after the measurement, the system is in the state corresponding to the state vector $|\psi_j\rangle$.

In case the state $\rho = |\psi\rangle\langle\psi|$ is pure (i.e., is a rank one projection and an extreme point of state space) one obtains

$$\mathbb{P}(j) = |\langle \psi_j | \psi \rangle|^2. \tag{17}$$

Let us have a look at two simple examples. The most prominent observable for a simple qubit is the measurement of a *Pauli operator* (or a Pauli matrix, if represented in a given basis). In fact, in a multi-qubit setting, they are again very important, where they give rise to a group, the *Pauli group* [44].

Definition 3 (Pauli matrices) *The Pauli matrices are given by*

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \mathbb{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{18}$$

They are unitary and Hermitian, and constitute an operator basis of the set of Hermitian matrices in $\mathbb{C}^{2\times 2}$.

They are Hilbert-Schmidt orthogonal,[2] as can easily be verified. These matrices are not only mathematically important, but also on physical grounds. The eigenbasis of Z is given by $\{|0\rangle, |1\rangle\}$ (taking this basis as the reference basis). The eigenbasis of X, in contrast, is given by

$$\{|+\rangle := (|0\rangle + |1\rangle)/\sqrt{2},\ |-\rangle := (|0\rangle - |1\rangle)/\sqrt{2}\}. \tag{19}$$

If one performs a Z measurement and the state is initially in the state vector $|0\rangle$, then the probability of obtaining the outcome labeled 0 is unity, $\mathbb{P}(0) = |\langle 0|0\rangle| = 1$ and the outcome labeled 1 is $\mathbb{P}(1) = |\langle 1|0\rangle| = 0$, so that one would receive the same outcome over and over again. However, if one performs on the same state an X measurement, one gets

$$\mathbb{P}(0) = |\langle + |0\rangle|^2 = \frac{1}{2},\ \mathbb{P}(1) = |\langle - |0\rangle|^2 = \frac{1}{2}. \tag{20}$$

That is to say, if the state vector was initially in an eigenstate of Z (with eigenvectors $|0\rangle, |1\rangle$) and one performs a measurement of the observable X (with eigenvectors $|+\rangle, |-\rangle$), both outcomes are equally probable. A repeated measurement of the observable X will then yield the same outcome over and over again.

[2] The *Hilbert-Schmidt scalar product* of A and B is given by $(A, B) := \text{tr}(A^\dagger B)$.

2.5 BB84 quantum key distribution scheme as an example

We now turn to our first cryptographic application. This scheme, the famous BB84 scheme for *quantum key distribution* [5], was the historically first scheme for quantum key distribution, featuring in its name the initial letters of Bennett and Brassard, the names of its inventors. It is built on earlier work by Wiesner on *quantum money* and *conjugate coding* [65]. It is both an ingenious scheme that lives up to the expectations of a modern quantum key distribution scheme and it also serves as a nice example of the quantum formalism laid out above.[3] It is based on the iterated use of single qubits only, so the state spaces we need to consider are merely $\mathcal{S}(\mathbb{C}^2)$. Interestingly, rigorous security proofs were only found more than a decade later [39, 57].

In the BB84 protocol, Alice sends a string of qubits to Bob, prepared one by one, and hence in a product state. She prepares either states that are eigenstates of Pauli Z or she prepares eigenstates of Pauli X. That is to say, she prepares orthogonal states taken from two non-orthogonal bases. Specifically, the protocol proceeds as follows.

- Alice picks an i.i.d. random bit string $a \in \{0, 1\}^n$.
- Alice picks a second i.i.d. random bit string $b \in \{0, 1\}^n$. At this point, she does not reveal either of the two bit strings.
- Alice now prepares quantum states of single qubits that she sends to Bob. The basis picked will depend on b: If $b_i = 0$, she prepares the i-th state in the Z basis with eigenvectors $\{|0\rangle, |1\rangle\}$, if $b_i = 1$, she prepares states in the X basis with eigenvectors $\{|+\rangle, |-\rangle\}$. That is to say, for the following values (b_j, a_j) she prepares

$$(0,0): \quad |0\rangle, \tag{21}$$
$$(0,1): \quad |1\rangle, \tag{22}$$
$$(1,0): \quad |+\rangle, \tag{23}$$
$$(1,1): \quad |-\rangle. \tag{24}$$

  Note that $|0\rangle, |1\rangle \in \mathbb{C}^2$ are orthogonal, and so are $|+\rangle, |-\rangle \in \mathbb{C}^2$, but the respective bases are not orthogonal to each other.
- These states are sent to Bob via a quantum channel.
- Bob picks an i.i.d. random bit string $c \in \{0, 1\}^n$.
- Depending on the value c_j, Bob measures in the Z basis ($c_j = 0$) or the X basis ($c_j = 1$).
- If the basis picked by Bob is the same one as the one Alice picked, so if for a value j one has $c_j = b_j$, then the outcome of the measurement will be deterministic: Bob will receive the measurement outcome d_j. If no eavesdropper is present, $d_j = a_j$ with certainty, following the above rule for quantum measurements.

[3] The recollection of how the BB84 scheme came about is an interesting story in its own right. Note also that Bennett, a theoretical physicist, actually implemented a first experimental demonstration of the scheme himself, using a setup that is still standing on his desk in his office at the IBM Watson Center.

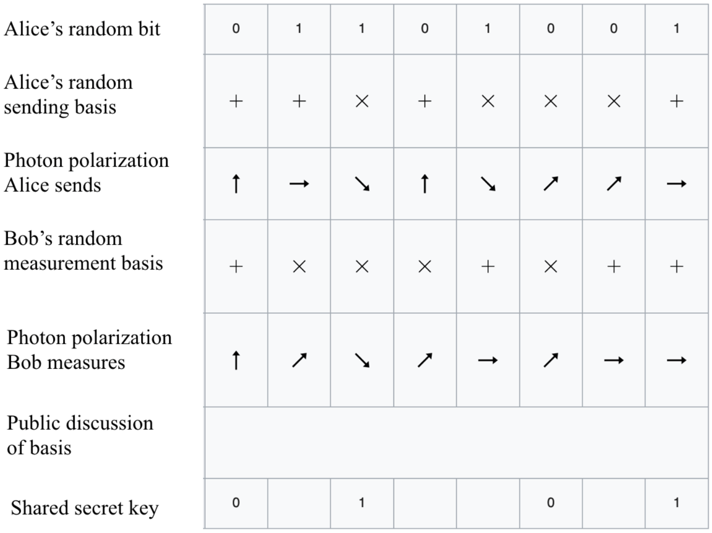
- If in contrast the basis picked is different, so if for a value j one has $c_j \neq b_j$, then he will receive an i.i.d. random number, not correlated with a_j. At this point, Bob cannot judge, however, which is the case, since at this point, Bob has not received any classical information from Alice yet.

- Alice and Bob communicate classically over the bases used, so they reveal the bit strings b and c. The string a is not revealed at any time, however.

- Alice and Bob discard all cases j for which $c_j \neq b_j$. This will happen in expectation in half the cases. They end up with a bit string I of expected length $n/2$.

- They take the measurement outcomes and values a_j, $j \in I$.

- In order to determine the presence of an eavesdropper, Alice and Bob now compare a predetermined subset $J \subset I$ of the bit string I established. According to the quantum mechanical rules, the values $d_j = a_j$ should follow, if no third party (an "eavesdropper", commonly referred to as Eve) was present. If an eavesdropper has gained any information about the quantum states sent, this must introduce errors in Bob's measurements. Other environmental conditions can give rise to errors of the same type. If the rate of bits differing in J is $p > p_0$, they will abort the key and try again, possibly with a different quantum channel, as the security of the key can not be guaranteed under these circumstances. The threshold value p_0 is chosen so that the number of bits available to the eavesdropper Eve is less than this number, privacy amplification can be used to reduce Eve's knowledge of the key to an arbitrarily small amount at the cost of reducing the length of the key.



Alice's random bit	0	1	1	0	1	0	0	1
Alice's random sending basis	+	+	×	+	×	×	×	+
Photon polarization Alice sends	↑	→	↘	↑	↘	↗	↗	→
Bob's random measurement basis	+	×	×	×	+	×	+	+
Photon polarization Bob measures	↑	↗	↘	↗	→	↗	→	→
Public discussion of basis								
Shared secret key	0		1			0		1

The remaining bit string $I \backslash J$ is the raw key. Why does this give rise to a secure key? We will look at security proofs later. The point is that by the time the quantum systems are being sent, Eve has no chance to guess the correct basis any better than making random choices. In this way, she has to introduce errors to the quantum state with high probability. In case she guessed right, she will get the right outcome, as $|\langle 0|0\rangle| = |\langle 1|1\rangle| = 1$. In the other cases, however, she will get uniformly random outcomes, as $|\langle +|0\rangle| = |\langle -|1\rangle| = 1/2$. Of course, she is not forced to precisely do such measurements, as she is free in her choices. However, this will not help her. This is a consequence of a very basic theorem we will encounter later. By the time the measurement bases are revealed, it is too late, and she cannot make use of that information any more. It is the key point of quantum key distribution that this idea does not only work if Eve sticks to performing von-Neumann measurements in the given basis. It works if Eve performs arbitrary measurements, even ones that are entangled over all invocations of the preparations, and making unrealistic assumptions about her. She might even have a quantum computer at her disposal, allowing for arbitrary coherent manipulation of all qubits sent. Still, asymptotically, she will not gain information about the key.

2.6 Entangled and separable quantum states

We will end this section by discussing *entangled states*. They would more naturally belong to the discussion of quantum state spaces, but for didactical reasons we have moved this subsection here. Entanglement is a property of bi-partite or multi-partite quantum systems. It reflects *quantum correlations* that are in a sense stronger than classical correlations in probability distributions. We label here the tensor factors A and B, in some connotation to Alice and Bob, as such states are usually considered in spatially distributed settings.

Definition 4 (Entangled and separable quantum states) *A state vector* $|\psi\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ *is called* entangled *iff it is not a product, i.e., if it cannot be written in the form*

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle \tag{25}$$

with $|\phi_A\rangle \in \mathbb{C}^{d_A}$ *and* $|\psi_B\rangle \in \mathbb{C}^{d_B}$. *A general quantum state is called* entangled *[64], iff it is not* classically correlated *or* separable*, i.e., contained in the* convex hull of products*, i.e., if it cannot be written as*

$$\rho = \sum_j p_j (\rho_A^j) \otimes (\rho_B^j), \tag{26}$$

with $\rho_A^j \in \mathcal{S}(\mathbb{C}^{d_A})$ *and* $\rho_B^j \in \mathcal{S}(\mathbb{C}^{d_B})$*, and* p *is a classical probability distribution. The* set of separable states $\mathcal{S}_{Sep}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}) \subset \mathcal{S}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ *is a convex subset of the set of separable states with a finite volume [35].*

An example of a product state vector of two qubits is

$$|\psi\rangle = |0,0\rangle. \tag{27}$$

Then each qubit is associated the state vector $|0\rangle$. A measurement on the respective subsystems will show no correlations whatsoever. The situation is entirely different for the state vector

$$|\psi\rangle = (|0,0\rangle + |1,1\rangle)/\sqrt{2}. \tag{28}$$

In fact, a Z measurement in both tensor factors A and B would lead to maximally correlated outcomes. The only outcomes one obtains are $(0,0)$ and $(1,1)$, but never $(0,1)$ or $(1,0)$. How can one find out that a given state vector is not a product? The partial trace is the key.

Definition 5 (Partial trace) *The reduced quantum state of a quantum state* $\rho \in \mathcal{S}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ *with respect to the first tensor factor is the unique quantum state* $\rho_A \in \mathcal{S}(\mathbb{C}^{d_A})$ *that satisfies*

$$\operatorname{tr}(\rho(A \otimes \mathbb{I})) = \operatorname{tr}(\rho_A A), \tag{29}$$

for all $A = A^\dagger$ *acting on* $\mathbb{C}^{d_A}$*. The linear map* $\rho \mapsto \rho_A$ *is called the partial trace.*

It is easy to see that a partial trace is indeed this, a "partial" trace, in that one picks a basis $\mathcal{B}_B$ of $\mathbb{C}^{d_B}$ and then performs the trace over this basis. Since it is not a complete basis of the full vector space $\mathcal{H} = \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, the remaining object is not a real number but a quantum state. It is not difficult to see (making use of the eigenvalue decomposition) that the choice of basis in $\mathbb{C}^{d_B}$ is irrelevant for the partial trace. The interpretation of the reduced quantum state is basically that of the marginal of a classical probability distribution.

There is something remarkable about the partial trace of quantum states, however. $(|0,0\rangle + |1,1\rangle)/\sqrt{2}$ is a state vector and hence

$$\rho = \frac{1}{2}(|0,0\rangle + |1,1\rangle)(\langle 0,0| + \langle 1,1|) \tag{30}$$

a pure state, an extreme point of state space. Its reduced state ρ_A obtained upon performing the partial trace

$$\rho_A = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \tag{31}$$

is not only not pure: Is it *maximally mixed*, in matrix form

$$\rho_A = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{32}$$

A quantum state is maximally mixed iff it is proportional to the identity. It reflects a uniform mixture in classical probability theory. So the *marginal* of a pure state is in general no longer pure.

2.7 Computational complexity of the separability problem

How can one determine whether a mixed quantum state is entangled or not? Slightly more formally, the problem at hand is the following membership problem. To make this a little bit more precise, let us define the following two sets: Let us first write $\mathcal{S} := \mathcal{S}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ and $\mathcal{S}_{\mathrm{Sep}} := \mathcal{S}_{\mathrm{Sep}}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ for brevity. We define

$$S(\mathcal{S}_{\mathrm{Sep}}, \delta) := \left\{\rho \in \mathcal{S} : \exists \sigma \in \mathcal{S}_{\mathrm{Sep}} \text{ with } \|\rho - \sigma\|_2 < \delta\right\} \tag{33}$$

as the "deeply separable states" up to an accuracy $\delta > 0$ and

$$S(\mathcal{S}_{\mathrm{Sep}}, -\delta) := \left\{\rho \in \mathcal{S}_{\mathrm{Sep}} : S(\rho, \delta) \subset \mathcal{S}_{\mathrm{Sep}}\right\}. \tag{34}$$

Then the separability problem becomes what is called a weak membership problem:

Definition 6 (Separability problem [36]) *Let* $\rho \in \mathcal{S}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}) = \mathcal{S}$ *and* $\delta > 0$ *be a a precision parameter.*

$$\begin{aligned} \textit{If } \rho \in S(\mathcal{S}_{Sep}, -\delta) \quad & \textit{output YES,} && (35) \\ \textit{if } \rho \notin S(\mathcal{S}_{Sep}, \delta) \quad & \textit{output NO.} && (36) \end{aligned}$$

Here $\|.\|_2$ is the *Frobenius norm*, defined as

$$\|A\|_2 = \mathrm{tr}(A^2) = (A, A) \tag{37}$$

for matrices A. However, from the perspective of computational complexity, this turns out to be a computationally hard problem.

Theorem 7 (Hardness of the separability problem [27]) *The separability problem is* NP*-hard, if* δ *scales inversely exponentially with respect to the dimensions* d_A, d_B.

The proof, laid out in Ref. [27], makes use of a polynomial-time reduction to Edmonds' problem. The original proof shows that this problem is NP-hard if ρ is located within an inverse exponential (with respect to dimension) distance from the border of the set of separable quantum states. In Ref. [25] the proof of NP-hardness is extended to an inverse polynomial distance from the separable set.

This does not mean, however, that one cannot find efficient one-sided tests. In particular, there exist *entanglement witnesses*, so observables $W = W^\dagger$ with the property that

$$\mathrm{tr}(W\rho) \geq 1 \ \forall \rho \in \mathcal{S}_{\mathrm{Sep}}. \tag{38}$$

and there exists an entangled state $\sigma \in \mathcal{S} \backslash \mathcal{S}_{\mathrm{Sep}}$ with

$$\mathrm{tr}(W\sigma) < 1. \tag{39}$$

From the perspective of convex analysis, this is a separating hyperplane [11] of the convex set of separable states $\mathcal{S}_{\mathrm{Sep}}$. It is an *optimal entanglement witness* [38] if it is a tangent hyperplane. Finding optimal witnesses translates to shifting a given entanglement witness to become an optimal entanglement witness and is again an NP-hard problem. In practice, entanglement witnesses are important: $\mathrm{tr}(W\rho)$ are expectation values of W that are accessible in experiments. Whenever one finds a value $\mathrm{tr}(W\sigma) < 1$ one can unambiguously argue that the unknown quantum state must have been entangled.

3 Quantum channels

3.1 Quantum channels and complete positivity

We have hinted at quantum key distribution schemes being secure against general strategies of eavesdropping, but have not clarified yet what such general strategies amount to. This relates to the question what the most general transformation is that one can apply to quantum states. This is given by the notion of a *quantum channel.* Mathematically speaking, quantum channels capture the legitimate transformations that quantum states can undergo. They directly generalize the concept of a *stochastic matrix* that maps probability vectors onto probability vectors. A stochastic matrix is a matrix $P \in \mathbb{R}_+^{d\times d}$ with the property that

$$\sum_{k=1}^{d} P_{j,k} = 1, \tag{40}$$

so that probability vectors $p \in \mathbb{R}_+^d$ with $\sum_{j=1}^d p_j = 1$ are mapped to probability vectors.

A quantum channel captures two things: First, it is the most general valid operation that one can perform on a quantum system. Second, it describes real communication channels as a special case. From the perspective of a mathematical characterization, what defines a quantum channel T? Surely, such channels must be linear maps, so that if T_1 and T_2 are quantum channels, then

$$T = \alpha T_1 + \beta T_2 \tag{41}$$

with $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$ is a quantum channel. A quantum channel T must also be a *positive map*, $T \geq 0$, mapping positive operators $\rho \geq 0$ onto positive operators, such that

$$\sigma = T(\rho) \geq 0 \tag{42}$$

must again be a valid positive operator. Interestingly, this turns out not to be enough: One needs a stronger form of positivity, referred to as *complete positivity*.

Definition 8 (Complete positivity and quantum channels) *Linear maps* T *on* $\mathcal{H}$ *are called completely positive iff*

$$T \otimes \mathrm{id} \geq 0, \tag{43}$$

where $T \otimes \mathrm{id}$ *is a linear map on* $\mathcal{H} \otimes \mathcal{H}$ *with* $\mathcal{H} = \mathbb{C}^d$. *Quantum channels are trace-preserving completely positive maps satisfying*

$$\mathrm{tr}(T(\rho)) = 1 \tag{44}$$

for all ρ *satisfying* $\mathrm{tr}(\rho) = 1$.

It turns out that it is sufficient to take d having the same dimension as the dimension of the first tensor factor. Why is that? Because T could act on a part of a larger system, and then the operator $(T \otimes \mathrm{id})(\rho)$ must again be a valid quantum state. This is a feature of quantum mechanics absent in classical mechanics: Although the map acts only on a part of the system and “does nothing” to the second tensor factor, the joint map still needs to be a positive map. The best known example of a positive but not completely positive map is the *transposition* t, mapping

$$t : \rho \mapsto \rho^T = \rho^*. \tag{45}$$

Note that ^{T} denotes the element-wise transposition and * the element-wise complex conjugation. Hermitian conjugation will be denoted by †. Physically, this map reflects a *time reversal.* It is easy to see that this is a positive map, so whenever $\rho \geq 0$ then also $\rho^T \geq 0$. But *partial transposition* is not completely positive. Think of the quantum state of two qubits in $\mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2)$, given by

$$\rho = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \tag{46}$$

with eigenvalues $\{1, 0, 0, 0\}$. Its partial transposition is then

$$\rho = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{47}$$

clearly not a positive matrix since it has eigenvalues $\left\{-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\}$. In fact, for single qubits, this is basically already all there is for positive matrices. We do not offer a proof of this statement.

Theorem 9 (Structure theorem for positive maps on qubits) *An arbitrary positive linear map* T *acting on* $\mathbb{C}^2$ *can be written as*

$$T = \alpha T_1 \circ t + \beta T_2 \tag{48}$$

with $\alpha, \beta \geq 0$*,* t *is the transposition and* T_1, T_2 *are completely positive linear maps.*

3.2 Choi-Jamiolkowski isomorphism and quantum channels as convex sets

More important is the following: The above definition of complete positivity does not give rise to a criterion that can be efficiently checked. Fortunately, the following statement provides such a criterion: It is necessary and sufficient for complete positivity to apply the linear map to a certain single reference state.

Theorem 10 (Criterion for complete positivity) *A linear map* T *on* $\mathcal{H}$ *is completely positive iff*

$$(T \otimes \mathrm{id})(\Omega) \geq 0 \tag{49}$$

where $\Omega \in \mathcal{H} \otimes \mathcal{H}$ *is a maximally entangled state.*

Proof: We will briefly prove this statement. We will need the following tiny Lemma for this: For any $\mathbb{C}^{d \times d} \ni P \geq 0$ and any $A \in \mathbb{C}^{d \times d}$, we have that

$$APA^\dagger \geq 0. \tag{50}$$

This is an immediate consequence of the fact that for every $|\psi\rangle \in \mathbb{C}^d$,

$$\langle\psi|APA^\dagger|\psi\rangle = (\langle\psi|A)P(A^\dagger|\psi\rangle) \geq 0. \tag{51}$$

Let us assume that Eq. (49) holds true. We will now show that

$$(T \otimes \mathrm{id})(\rho) \geq 0 \tag{52}$$

for all $\rho \in \mathcal{S}(\mathcal{H} \otimes \mathcal{H})$. We make use of the spectral decomposition

$$\rho = \sum_{j=1}^{d^2} p_j |\psi_j\rangle\langle\psi_j|. \tag{53}$$

From linearity, we have

$$(T \otimes \mathrm{id})(\rho) = \sum_{j=1}^{d^2} p_j (T \otimes \mathrm{id})(|\psi_j\rangle\langle\psi_j|), \tag{54}$$

so it is sufficient to show the statement for single state vectors $|\psi\rangle \in \mathcal{H}$. Now every such state vector can be written as

$$|\psi\rangle = (\mathbb{I} \otimes X)|\Omega\rangle \tag{55}$$

for a suitable $X \in \mathbb{C}^{d \times d}$. But then we have

$$\begin{aligned}(T \otimes \mathrm{id})(|\psi\rangle\langle\psi|) = (T \otimes \mathrm{id}) &\left((\mathbb{I} \otimes X)|\Omega\rangle\langle\Omega|(\mathbb{I} \otimes X^\dagger)\right) \\ &(\mathbb{I} \otimes X)(T \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|)(\mathbb{I} \otimes X^\dagger) \geq 0 \end{aligned} \tag{56}$$

from which the statement follows. In fact, it turns out that $(T \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|)$ completely specifies the channel.

Theorem 11 (Choi-Jamiolkowski isomorphism) *Quantum channels as completely positive, trace preserving maps* T *on* $\mathcal{H}$ *are isomorphic to the quantum states*

$$(T \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|). \tag{57}$$

The proof is left as an exercise. In fact, one direction of the proof of the isomorphism we have already elaborated upon. This may not be a particularly deep statement, but it has profound implications. Channels can be viewed as quantum states on a larger Hilbert space. That also comes along with the insight that the set of quantum channels is again a convex set. In fact, any kind of optimization of linear functionals over quantum channels can be cast into the form of a *convex optimization problem*. We will see that in fact *semi-definite programming* [11] is at the heart of the optimization of many quantum protocols.[4] In fact, many optimal success probabilities of protocols can readily be captured as semi-

[4] Semi-definite programming [11] generalizes linear programming and is a form of a convex optimization problem for which the theory is very much developed, and for which interior point methods provide an efficient solution. They are optimization problems of the form, for vectors $c \in \mathbb{R}^d$ and matrices $F_0, \dots, F_d \in \mathbb{R}^{D \times D}$

$$\text{minimize } c^T x, \tag{58}$$

$$\text{subject to } F_0 + \sum_{j=1}^{d} x_j F_j \geq 0. \tag{59}$$

The Lagrange dual is again a semi-definite problem of the form

$$\begin{aligned} \text{maximize } & -\mathrm{tr}(Z F_0), & (60)\\ \text{subject to } & \mathrm{tr}(Z F_j) = c_j \forall j = 1, \dots, d, & (61)\\ & Z \geq 0. & (62)\end{aligned}$$

Any solution to the Lagrange dual provides a lower bound to any solution to the original, the primal, problem, which is a property most useful when using semi-definite programming in proofs.

definite programs of this form [3]. In a bigger picture, ideas of convex [4] and non-convex programming [19, 21] feature strongly in quantum mechanics. In fact, the latter works provide hierarchies of semi-definite programs to decide the above separability problem, where each level of the hierarchy can be solved in polynomial time.

3.3 Kraus' theorem and Stinespring dilations

We have understood what a completely positive map is, but not how it can be parametrized and what specific form it takes. This is given by Kraus' theorem.

Theorem 12 (Kraus' theorem) *A linear map* T *on* $\mathcal{H}$ *is completely positive and trace-preserving exactly if it can be written as*

$$\mathrm{T}(\rho) = \sum_{j=1}^{r} K_j \rho K_j^\dagger \tag{63}$$

satisfying

$$\mathrm{T}(\rho) = \sum_{j=1}^{r} K_j^\dagger K_j = \mathbb{I}. \tag{64}$$

The smallest number r *that can be achieved in such a decomposition is called the Kraus rank.*

We do not have the time to present the full proof of this. But we sketch the idea. One direction of the proof is trivial: T is linear by construction. Also, applying (71) to (49) immediately gives rise to a positive operator. The more technical direction is to show that such a form can always be achieved. The key steps are to start from the spectral decomposition

$$(\mathrm{T} \otimes \mathrm{id})(\Omega) = \sum_i p_i |e_i\rangle\langle e_i|. \tag{65}$$

Now take an arbitrary state vector $|\psi\rangle \in \mathcal{H}$, and to extend it onto $\mathcal{H} \otimes \mathcal{H}$ as $|\psi^*\rangle \otimes |\psi\rangle$. One can then write

$$|\psi\rangle\langle\psi| = d\langle\psi^*|\Omega\rangle\langle\Omega|\psi^*\rangle = d\mathrm{tr}_A(|\psi^*\rangle\langle\psi^*| \otimes \mathbb{I})|\Omega\rangle\langle\Omega|). \tag{66}$$

Then applying T can be done on the second tensor factor. The Kraus operators are then defined by

$$K_j|\psi\rangle = \sqrt{dp_i}\langle\psi^*|e_i\rangle. \tag{67}$$

Note that the Kraus decomposition is not unique: Any set $\{l_k\}$ is again a set of Kraus operators if

$$l_k = \sum_i U_{k,i} K_i \tag{68}$$

for U being unitary is again a legitimate set of Kraus operators. It is also not difficult to see that the Kraus rank is exactly the standard rank of the Choi-Jamiolkowski isomorph $(\mathrm{T} \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|)$, an insight that is again left to the reader as an exercise.

Any channel can be seen as a unitary map in a larger vector space, a statement captured by *Stinespring's theorem*. We will spell it out in a slightly unusual and redundant form, yet one that is easier to communicate. This is a most important form: Its significance stems from the observation that unitary operations originate from time evolution in quantum mechanics, the most important quantum channel.

Definition 13 (Hamiltonian evolution) *The channel*

$$\rho \mapsto U\rho U^\dagger \tag{69}$$

with U *being a unitary on* $\mathcal{H}$ *captures* Hamiltonian time evolution *generated by a* Hamiltonian $H = H^\dagger$ *via* $U = \exp(-itH)$. *Such dynamics is referred to as* Schrödinger dynamics.

In fact, most elementary courses on quantum mechanics elaborate on the consequences of such time evolution generated by meaningful Hamiltonians capturing important physical systems: The Schödinger equation is one of the key equations and one of the axioms of quantum mechanics. The point of the Stinespring dilation is now to see that any channel can be seen as such a unitary channel on a larger vector space.

Theorem 14 (Stinepring dilations) *Any completely positive and trace-preserving map* T *on* $\mathcal{H} = \mathbb{C}^d$ *can be written as*

$$T(\rho) = \mathrm{tr}_2(U(\rho \otimes \eta)U^\dagger) \tag{70}$$

where η *is a quantum state on* $\mathbb{C}^D$, U *is a unitary defined on* $\mathbb{C}^d \otimes \mathbb{C}^D$, *and* tr_2 *is the partial trace with respect to the second tensor factor.* D *has at most be taken to be* d.

3.4 Disturbance versus information gain

This insight also clarifies what the most general attack in the above eavesdropping scheme is: Any eavesdropper can take a further system initially in a state η and entangle it with the system at hand in state ρ. Then she can perform measurements on her system. In fact, the labels of the Kraus' theorem exactly correspond to the labels in a von-Neumann measurement when the measurement postulate is applied to the system initially in η.

Theorem 15 (Generalized measurement) *The Kraus decomposition can be realized as*

$$K_j \rho K_j^\dagger = \mathrm{tr}_2((\mathbb{I} \otimes \pi_j)U(\rho \otimes \eta)U^\dagger) \tag{71}$$

where ω *is a quantum state on* $\mathbb{C}^D$, U *is a unitary defined on* $\mathbb{C}^d \otimes \mathbb{C}^D$, tr_2 *is the partial trace with respect to the second tensor factor, and* $\pi_j = |\psi_j\rangle\langle\psi_j|$ *are unit rank projections from the measurement postulate.*

We will now go too much into detail here: But when captured in this form, it should be clear that the information gain (the knowledge obtained via the statistics of measurement outcomes) and the disturbance (the alteration of ρ to the state conditioned on measurement outcomes) are in a close relationship to one another. The *disturbance versus information* gain has first been comprehensively studied in Ref. [24].

4 Channel capacities and quantum information transmission

The notion of a channel capacity captures what rate of information can be transmitted via a given communication channel, let this be quantum or classical. In the context of quantum communication, naturally, quantum channels are in the focus of attention.

4.1 Diamond norms

When considering channel capacities for quantum channels, we first need to know in what sense we can approximate a quantum channel. A starting point is the *trace-norm distance* for quantum states: It is defined for two quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ as

$$D(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1 = \frac{1}{2}\mathrm{tr}|\rho - \sigma|, \tag{72}$$

where $|\cdot|$ denotes the operator absolute value (so the sum of the singular values of the argument). A moment of thought reveals (by invoking the above Kraus theorem) that it quantifies the statistical distinguishability of ρ from σ. This distance hence operationally captures how different ρ is from σ. More specifically, $\frac{1}{2}(1 + D(\rho, \sigma))$ is the maximal success probability when trying to distinguish ρ and σ via measurement. Let us now move to meaningful distance measures for quantum channels. One might think a good distance measure for two quantum channels T and S on $\mathcal{H} = \mathbb{C}_d$ (i.e., completely positive, trace-preserving maps) is

$$\|T - S\|_1 = \sup_{\|A\|_1 = 1} \|T(A) - S(A)\|_1. \tag{73}$$

However, there is a problem with this definition: The norms of $\|T \otimes \mathrm{id}_n\|_1$ may increase with n, even though the channel does not even act non-trivially on the second tensor factor. For this reason, one defines the *diamond norm* distance (the factor $1/2$ has no signifiance) as follows.

Definition 16 (Diamond norm) *For two channels T and S, the diamond norm is defined as*

$$\|T - S\|_\diamond = \sup_n \sup_{\|A\|_1 = 1} \|(T \otimes \mathrm{id}_n)(A) - (S \otimes \mathrm{id}_n)(A)\|_1. \tag{74}$$

Given the clumsy definition as an unbounded supremum, one might be tempted to think that this norm cannot be computed. In fact, it can, even efficiently: It turns out to be the solution, once again, of a semi-definite problem [63].

4.2 Capacities as asymptotic transmission rates

The notion of a capacity asks at what rate information can be transmitted. Here, "rate" refers to an asymptotic rate, invoking a communication channel more than once. This makes a lot of sense, and classical channel capacities are also defined in such a fashion. We can define capacities of quantum channels T based on the quantity

$$\Delta(S, T) = \inf_{D,E} \|S - D \circ T \circ E\|_\diamond, \tag{75}$$

as the infimum over encoding channels E and decoding channels D . After all, we are only interested in the optimal encoding and decoding. S here is the identity channel over a certain algebra. S is seen as representing a word of the kind of message that is supposed to be sent, whereas T stands for a single invocation of the channel. However, when defining a capacity, we are less interested in single invocations, but rather in many invocations and long messages. This refers to the situation of considering $T^{\otimes n}$ and $S^{\otimes n}$ for large n, but encodings and decodings over many channels of this type.

Definition 17 (Capacity) *Let S and T be quantum channels. Then a number $c \geq 0$ is called an "achievable rate" for T with respect to S, if for any sequences n_α, m_α of integers with $m_\alpha \to \infty$ and*

$$\limsup_\alpha \left(\frac{n_\alpha}{m_\alpha}\right) < c \tag{76}$$

we have

$$\lim_\alpha \Delta(S^{\otimes n_\alpha}, T^{\otimes m_\alpha}) = 0. \tag{77}$$

The supremum of all achievable rates is called the capacity *of T with respect to S and is denoted by* $C(S, T)$.

4.3 Classical information capacity and additivity problems

The classical information capacity C_c is defined as the rate at which classical bits can be sent via a quantum channel. The quantum capacity C_q in turn is the rate at which quantum bits, qubits, can be transmitted. If the one-bit system is defined as $\mathcal{C}_2$ and the one qubit system as $\mathcal{M}_2$, then they are

$$\begin{aligned} C_c(T) &= C(\mathcal{C}_2, T), && (78) \\ C_q(T) &= C(\mathcal{M}_2, T). && (79) \end{aligned}$$

It is clear that

$$C_q(T) \leq C_c(T) \tag{80}$$

for any quantum channel T, as quantum channels can be used to send classical bits. But it may be true that some noisy channels only allow for the transmission of classical information, but no coherent quantum information. These capacities are notoriously difficult to compute. However, stringent bounds can be found, and formulae do exist. The most famous result is the expression of the classical information capacity [33, 34, 52]. In order to state this, we need to define the *von-Neumann entropy* of a quantum state. It is the quantum analogon of the Shannon entropy and given by

$$S(\rho) = -\text{tr}(\rho \log \rho), \tag{81}$$

in terms of a matrix logarithm (that is computed on the spectrum of ρ as a matrix function). The classical information capacity is then found to be the following expression.

Definition 18 (Classical information capacity) *The single-shot classical capacity is given by*

$$C_{c,1}(T) := \max_{\{p_i,\rho_i\}} \left(S(\sum_i p_i T(\rho_i)) - \sum_i p_i S(T(\rho_i)) \right), \tag{82}$$

where $\{p_i\}$ is a probability distribution and $\{\rho_i\}$ is a set of quantum states. The actual classical information capacity is then regularized as

$$C_c(T) := \sup_n \frac{1}{n} C_{c,1}(T^{\otimes n}). \tag{83}$$

This expression seems puzzling: How can it be an advantage to send information coherently over many channels, so why is not simply $C_c(T) = C_{c,1}(T)$? It turns out that it does help. It was an open problem for a long time whether or not the classical information capacity was *additive*. Using ingenious ideas of random coding, it was shown in large dimension to be beneficial to use entangled inputs, even though only classical information is to be transmitted. In fact, the additivity was a long-standing puzzle and open question in the field: It could be shown that many additivity questions in quantum information theory were equivalent [2, 56], until it was finally settled [29]. In fact, the latter publication made use of the proven equivalences and proved a counterexample for the additivity of the minimum output entropy of a quantum channel, maybe the most puzzling of the known additivity problems. It was done using random quantum coding, and it is still an open problem in the field to provide a constructive counterexample.

4.4 Quantum capacity and super-activation

How about the quantum capacity? For this, we define the *coherent information* as

$$C_{q,1}(T) := \sup_\rho \left(S(\rho_B) - S(\rho_E)\right), \tag{84}$$

where in a Stinespring dilation we write the channel as $T(\rho) = U(\rho \otimes \omega)U^\dagger$ and consider the output state as a bi-partite state over B and E. The genuine quantum capacity is again seen as an asymptotic limit.

Definition 19 (Quantum information capacity) *The quantum capacity is given by*

$$C_q(T) := \sup_n \frac{1}{n} C_{q,1}(T^{\otimes n}). \tag{85}$$

Again, it is known that the "regularization" on the right hand side is needed. This renders the quantum capacity a quantity that in practice cannot be computed, but bounded.

5 Quantum repeaters for secure long-distance quantum key distribution

5.1 Entanglement based key distribution schemes

In Section 2.5 we have encountered the BB84 scheme as a scheme for quantum key distribution in which quantum systems are being prepared and then sent through a quantum channel. It is the most important and still most practical scheme for quantum key distribution. Having said that, there are many other schemes for quantum key distribution. One way of categorizing them is into *prepare-and-measure* schemes (such as the original BB84 scheme) and into *entanglement-based schemes*. The latter type of scheme at first sight seems quite distinctly different: One first prepares an entangled state, to then – once distributed – performs local measurements to establish a key. However, a moment of thought reveals that this is something very similar: In fact, every prepare-and-measure scheme can be seen as an equivalent to an entanglement based scheme. Take the maximally entangled state vector of two qubits

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0,1\rangle - |1,0\rangle). \tag{86}$$

If Alice on one side performs a Z measurement and obtains 0 as her outcome, Bob's system will be in $|1\rangle$. Similarly, upon a 1 outcome, Bob's system will projected to be in $|0\rangle$. That is to say, the measurement on Alice's side is effectively like a (non-deterministic) preparation of a quantum state on Bob's side. Since the state vector is UU-invariant, i.e., $(U \otimes U)|\psi\rangle = |\psi\rangle$ for all $U \in U(2)$, the same holds true for X measurements, and in fact any other measurement in the same basis on both sides. If Alice projects her system into $|+\rangle$, Bob will see $|-\rangle$, and if she encounters $|-\rangle$, Bob will have $|+\rangle$. Again, this can be seen as a probabilistic preparation. This connection between prepare-and-measure schemes has long been observed. In fact, a precondition for security in any scheme, including prepare-and-measure schemes, is the presence of entanglement in the equivalent entanglement (or correlation) based scheme [13].

5.2 Entanglement swapping and distillation

However, there is an important conceptual difference: In prepare-and-measure schemes, there is little one can do about losses when sending quantum systems through quantum channels. In entanglement based schemes, one can do something about it. Accepting this, the key question is: How can one establish a maximally entangled state between arbitrary distances in the first place? This is possible by means of *quantum repeaters*. They consist of two steps:

- First, they involve *entanglement distillation*. Imagine two parties prepare maximally entangled states Ω and send one half through a quantum channel. Due to losses, they are being transformed into states

  $$\rho = (\mathrm{id} \otimes T)(\Omega), \tag{87}$$

  still entangled quantum states, but ones that are mixed and no longer maximally entangled. Upon n_α invocations of such preparations, the state prepared is $\rho^{\otimes n_\alpha}$, so n_α "copies" of the state ρ. One can now perform local operations, coordinated by classical communication, to transform $\rho^{\otimes n_\alpha}$ into $m(n)$ copies of approximately maximally entangled states Ω. These local operations will involve Kraus operators of the form

$$\{A_j \otimes \mathbb{I}\} \tag{88}$$

when implemented by Alice and

$$\{\mathbb{I} \otimes B_j\} \tag{89}$$

when implemented by Bob. The measurement outcomes can in all instances be communicated in protocols involving an arbitrary number of rounds. This will at best be possible at a rate

$$\limsup_{\alpha \to \infty} \frac{n_\alpha}{m_\alpha} < 1. \tag{90}$$

In effect, both parties will end up with fewer, almost maximally entangled states. Entanglement has been "distilled", in a similar way as one can extract high percentage alcohol from a liquid in which alcohol is only present in a dilute form. These maximally entangled states can be used in subsequent steps. This is a highly interesting procedure: Entanglement, so intrinsic quantum correlations, are here manipulated like an interconvertible resource.

- Then there are steps of *entanglement swapping*. This step is maybe even more intricate and interesting. Think of two maximally entangled states shared by Alice and Bob on the one hand and Bob and Charlie on the other hand. So let us start from

$$|\psi\rangle = |\Omega\rangle_{A,B_1} \otimes |\Omega\rangle_{B_2,C}, \tag{91}$$

with again

$$|\Omega\rangle = (|0,0\rangle + |1,1\rangle)/\sqrt{2}. \tag{92}$$

That is to say, Bob holds two halves of maximally entangled states. The two copies will not have any shared history. Now Bob can perform a projective measurement, projecting the state vector into

$$(\mathbb{I}_A \otimes \langle\Omega|_{B_1,B_2} \otimes \mathbb{I}_C)|\psi\rangle = \frac{1}{\sqrt{2}}|\Omega\rangle_{A,C}. \tag{93}$$

That is to say, after the projective measurement, A and C are in a maximally entangled state, even though these particles have no joint history whatsoever. The entanglement has been "swapped". Of course, in a projective measurement, this would only work in a probabilistic fashion. However, a moment of thought reveals that this can be made deterministic, in that for each outcome of a joint measurement on B_1 and B_2, one can find a Pauli correction on A and C so that deterministically, $|\Omega\rangle_{A,C}$ is reached. The reason for this is ultimately that

$$\{(\mathbb{I} \otimes \mathbb{I})|\Omega\rangle,\ (X \otimes \mathbb{I})|\Omega\rangle,\ (Y \otimes \mathbb{I})|\Omega\rangle,\ (Z \otimes \mathbb{I})|\Omega\rangle\} \tag{94}$$

for Pauli operators $X, Y, Z, \mathbb{I}$ constitute a basis of the maximally entangled states on $\mathbb{C}^2 \otimes \mathbb{C}^2$. The details of this, also constituting the basis of the very much related scheme of *quantum teleportation* [6] (for a review, see Ref. [47]), will be explained in the project period.

5.3 Full quantum repeater schemes

A *quantum repeater* now makes use of such such steps in a hierarchical, tree-like fashion. It involves steps of entanglement distillation between neighbours, followed by entanglement swapping steps. There are many variants of such quantum repeaters, as well as numerous suggestions for experimental realizations thereof. In fact, the reliable realization of quantum repeaters is the key obstacle on the path to secure quantum key distribution over arbitrary distances – while near-distance quantum key distribution is perfectly feasible. The trouble is that notions of entanglement distillation and entanglement swapping are generally assumed to rely on *quantum memories* that store quantum information reliably. Since quantum information has to be transmitted via light (other quantum systems are hardly feasible for this task), and quantum information is stored in matter qubits, one needs coherent frequency converters (to align the respective frequencies) and needs to map quantum states of light onto atoms, ions, or atomic ensembles. Then, it has to be read out, but needless to say, all coherently again at negligible losses. This still constitutes a technological road block, even though progress is fast. In fact, surprising as this may sound, each of the above mentioned components has already been achieved in experiments. Once this road block is overcome, secure quantum communication over arbitrary distances is feasible.

6 Quantum networks, graph theory, and routing

Quantum repeaters, once realized, enable the establishment of close to maximally entangled states over arbitrary distances. This allows for point-to-point quantum key distribution. But how can one think of multi-partite schemes and genuine *quantum networks*, involving a large number of nodes, and giving rise to genuine multi-partite schemes? This is what this section addresses, in a paradigmatic setting that largely abstracts from aspects of implementation. Recent years have, however, seen important progress in this respect []. Still, some aspects of genuine quantum networks can be captured in this manner.

6.1 Graph states

Any bi-partite state the reduction of which is maximally mixed is a maximally entangled state. This is true, e.g., for the pure state described on subsystems labeled 1 and 2 (and not A and B for reasons that will become clear later) by the state vector

$$|\psi\rangle = |0,0\rangle_{1,2} + |0,1\rangle_{1,2} + |1,0\rangle_{1,2} - |1,1\rangle_{1,2}. \tag{95}$$

One might think of this state being captured by two vertices connected by an edge that reflects the maximal entanglement. In fact, this state vector can be written as

$$U(|+\rangle_1 \otimes |+\rangle_2), \tag{96}$$

with U being a *controlled phase* quantum gate

$$U = |0,0\rangle\langle 0,0| + |0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0| - |1,1\rangle\langle 1,1|. \tag{97}$$

This unitary can be seen as entangling the two quantum systems initially prepared in

$$|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}. \tag{98}$$

Such an idea makes sense, in fact, for a quantum system composed of an arbitrary number of constituents. The idea of a *graph state* [30, 31, 50] captures very natural classes of *multipartite* quantum states [62].

We identify the nodes of a graph state with the vertices of a graph. An undirected, finite *graph* is a pair $G = (V, E)$ of a finite set $V \subset \mathbb{N}$ and a set $E \subset [V]^2$, the elements of which are subsets of V with two elements each [31]. The elements of V are *vertices*, the elements of E *edges*. When two vertices $a, b \in V$ are the endpoints of an edge, they are referred to as being *adjacent*. The adjacency relation gives rise to an *adjacency matrix* Γ_G associated with a graph. If $V = \{a_1, \dots, a_N\}$, then Γ_G is a symmetric $N \times N$-matrix, with elements

$$(\Gamma_G)_{i,j} = \begin{cases} 1, & \text{if } \{a_i, a_j\} \in E, \\ 0 & \text{otherwise.} \end{cases} \tag{99}$$

In what follows the neighbourhood of a given vertex $a \in V$ is important. This *neighbourhood* $N_a \subset V$ is defined as the set of vertices b for which $\{a, b\} \in E$. When the vertex a is deleted, together with the edges that are incident with a, the new graph obtained is called $G - \{a\}$. We now turn to describing how a given graph can be associated with a quantum state.

Definition 20 (Graph states) *For a given graph* $G = (V, E)$*, an associated* graph state *vector is now obtained by applying a sequence of controlled phase gates* $U^{(a,b)}$ *to the state vector* $|+\rangle^{\otimes|V|}$ *corresponding to the empty graph,*

$$|G\rangle = \prod_{(a,b)\in E} U_{a,b}\, |+\rangle^{\otimes|V|}, \tag{100}$$

where E *denotes the set of edges in* G*.*

Certain maximally entangled states of two qubits are graph states. Then, the *GHZ state* (Greenberger-Horne-Zeilinger-state) is also a graph state, in fact, one that is associated with the complete graph. Graph states are instances of *stabilizer states* that play an important role in quantum error correction. To every vertex $a \in V$ of $G = (V, E)$, one assigns the Hermitian operator

$$K_G^{(a)} = X_a \prod_{b\in N_a} Z_b \tag{101}$$

in terms of Pauli operators X and Z. Then the graph state vector $|G\rangle$ is the common eigenvector of the $K_G^{(a)}$ with unit eigenvalues, i.e.,

$$K_G^{(a)}|G\rangle = |G\rangle, \tag{102}$$

for all $a \in V$. The theory of quantum error correction is largely based on stabilizer states, for good reasons. It is also an efficiently describable class of pure quantum states: Instead of having to specify the coefficients of a state vector in an exponentially large Hilbert space, one merely has to state the polynomially many *Pauli operators* of which the state is an eigenvalue.

Definition 21 (Pauli group) *The Pauli group $\mathcal{G}_1$ on one qubit is the 16 element group defined by the Pauli operators together with prefactors ± 1 and $\pm i$, i.e.,*

$$\mathcal{G}_1 = \{\pm\mathbb{I}, \pm i\mathbb{I}, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}, \tag{103}$$

The Pauli group on n qubits $\mathcal{G}_n$ is the group generated by the operators described above applied to each of qubits in $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$.

Related to the Pauli group is the normalizer of the Pauli group, the so-called *Clifford group*.

Definition 22 (Clifford group) *The n-qubit Clifford group $\mathcal{C}_n$ is the normalizer (up to complex phases) of the Pauli group $\mathcal{G}_n$, i.e.,*

$$\mathcal{C}_n = \{U : U\mathcal{G}_n U^\dagger \subset \mathcal{G}_n\}/U(1). \tag{104}$$

Clifford operations, so the operations associated with Clifford unitaries, are extremely important in quantum information theory. Since they map Pauli operators onto Pauli operators, they can be efficiently described, even though they can be used to generate multipartite entanglement.

6.2 Local equivalence and local complementation

So far, the graph relation was rather a graphical representation than that graph theory really enters here. However, the connection between multipartite quantum states as they are relevant for quantum networks and graphs runs quite deep. Here we examine an important example of this type. If two graph states are the same up to local basis changes, they should be treated as being equivalent. In particular, local Clifford operations map graph states onto graph states. The respective graphs will, however, be different. How are they precisely related? In what follows, by $\tau_a(G)$ we denote the graph that results from locally complementing G with respect to the vertex a.

Definition 23 (Local complementation) *A graph $G = (V, E)$ and vertex $a \in V$ define a graph $\tau_a(G)$ having the adjacency matrix*

$$\Gamma_{\tau_a(G)} := \Gamma_G + \Theta_a \mod 2, \tag{105}$$

where Θ_a is the complete graph of the neighbourhood N_a.

The graph state that results from local complementation with respect to the vertex a of the graph state vector $|G\rangle$ [30, 31], is defined by $|\tau_a(G)\rangle = U_a^\tau|G\rangle$, where

$$U_a^\tau := (iX_a)^{1/2}(-iZ_{N_a})^{1/2}. \tag{106}$$

Needless to say, one can think of entire sequences of such local Clifford operations, and consider the entire orbit generated by such local graph complementations [59, 60]. Given that many graph problems are NP-hard, one might be tempted to think that this is a computationally hard task. Indeed, even the decision problem asking whether two graphs are equivalent up to a relabeling of the vertices is in NP. However, interestingly, one can assess whether one graph state can be obtained from another with a sequence of local complementations in polynomial time.

Theorem 24 (Bouchet theorem [10, 59]) *Given two graphs* $G = (V, E)$ *and* $G' = (V, E')$*, one can determine in time polynomial in* $|V|$ *whether* G *can be transformed into* G' *by a sequence of local complementations.*

As a corollary, one can determine in polynomial time whether two graph states are identical up to a local Clifford basis change that is of little physical significance [59].

6.3 Vertex minors of graphs and the quantum routing problem

Imagine that some source probabilistically prepares certain resource states. In some level of abstraction, these resource states can be seen as being maximally entangled after entanglement distillation steps. Given that the routing problem for quantum networks is still largely open, it seems reasonably to make use of that level of abstraction to make progress. Let us assume we would like to arrive at some desired target state. How can this target state be reached by means of a sequence of meaningful transformations? Let us assume that the initial state is given as a graph state vector $|G\rangle$. Can we bring this resource into the form, say, of a GHZ state, which is again a graph state? Meaningful operations are local Clifford basis changes and deletions of vertices as well as all edges that are incident to the deleted vertices. Physically, this corresponds to Pauli Z measurements on vertices. Mathematically, this gives rise to the vertex minor problem.

Theorem 25 (Vertex minor problem) *For a graph state associated with* G*, decide whether a graph* H *with* $|H| < |G|$ *can be extracted following a sequence of (i) local complementations and (ii) deletions of vertices. A graph* H *that can be obtained in this fashion is called a vertex-minor of* G*.*

Interestingly, the general vertex minor problem is a computationally hard problem.

Theorem 26 (Hardness of the vertex minor problem [16, 28]) *The vertex minor problem is* NP*-hard.*

For graphs with finite rank-width[5], one can, however, decide the routing problem in polynomial time.

Theorem 27 (Extraction of graph states from graph states with bounded rank-width) *For a graph state vector* $|G\rangle$ *with an underlying graph of bounded rank-width, there exists a poly-time algorithm that decides if a graph state vector* $|H\rangle$ *can be extracted from* $|G\rangle$ *using local Clifford operations and* Z*-measurements, and gives the sequence of operations to be applied.*

For important classes of target states, such as tripartite GHZ states, one can do the routing efficiently.

Theorem 28 (Extraction of tripartite GHZ states) *One can always extract a tri-partite GHZ state between arbitrary vertices* $a, b, c \in V$ *of a graph state vector* $|G\rangle$ *associated with a connected graph in polynomial time.*

Such tripartite GHZ states can be used in instances of quantum cryptographic protocols, specifically ones that go beyond bipartite quantum key distribution.

6.4 Quantum secret sharing as a multipartite cryptographic primitive

Multipartite states can be practically useful, for example, in *quantum secret sharing* schemes. This is a scheme for splitting a message into several parts so that no subset of parts is sufficient to read the message, but the entire set is. In classical secret sharing [53], one divides a secret into n shares such that at least t of those shares can be used to reconstruct the secret, while any $t-1$ or fewer shares have no information about the secret at all. Such a scheme is called an (t, n) threshold scheme. Quantum analogs of this idea are *quantum secret sharing* protocols [12, 32]. For brevity, we only give a particularly simple example of a $(2, 3)$ quantum threshold scheme, in which the constituents are not qubits, but qutrits, so that $\mathcal{H} = \mathbb{C}^3$. The encoding maps the secret qutrit for $\alpha, \beta, \gamma \in \mathbb{C}$ to three qutrits as

$$\begin{aligned} \alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle \quad &\mapsto \quad \alpha(|0,0,0\rangle + |1,1,1\rangle + |2,2,2\rangle) \\ &+ \quad \beta(|0,1,2\rangle + |1,2,0\rangle + |2,0,1\rangle) \\ &+ \quad \gamma(|0,2,1\rangle + |1,0,2\rangle + |2,1,0\rangle). \end{aligned} \tag{107}$$

[5] Rank width footnote. The *rank-width* k of a graph G is the minimum width of all its rank decompositions. This amounts to k being the smallest integer such that G can be related to a tree-like structure by recursively splitting its vertex set so that each cut induces a matrix of rank at most k. The rank-width is bounded iff the clique-width is bounded [46]. Graphs with rank-width at most one are those where all connected induced subgraphs preserve distance [45].

It is easy to see that no party alone can learn anything about the secret: Each reduced state is the maximally mixed state. But any two parties together can fully recover the state: Say, for the first two parties, one has to add the value of the first share to the second (modulo three), and then add the value of the second share to the first. Quantum secret sharing is a cryptographic primitive based on multi-partite states beyond bipartite settings.

7 Security proofs

7.1 Notions of attacks in quantum key distribution

Quantum key distribution offers the claim of secure key distribution in the presence of an eavesdropped that is attributed unlimited, even unrealistic, resources. Historically, however, before full security proofs were available, particular kinds of attacks were considered.

- *Intercept-resend attack:* The simplest type of a possible attack is the intercept-resend attack, in which Eve performs projective measurements, makes use of the measurement outcome, and prepares a new quantum system in a suitable state. It is easy to see that the BB84 scheme is secure against such intercept-resend attacks.
- *Individual attacks:* In this type of attack, Eve performs generalized measurements as laid out above, but interacts with each qubit (or other quantum system) in the channel separately and independently. Invoking the above Stinespring dilation, physically, this means Eve lets the quantum system transmitted interact with an auxiliary system each which is subsequently measured in a von-Neumann measurement. The intercept-resend attack is an instance of such an attack. Generally, individual attacks are the most realistic ones given present technology. *Photon number splitting attacks* in quantum optical schemes in which weak pulses are being sent are specifically important instances of such individual attacks.
- *Collective attacks:* This is a yet more general kind of attack. Here, Eve again performs generalized measurement. Again, she prepares independent auxiliary systems which interact with the quantum systems transmitted. But now she can perform a joint measurement on the collection of auxiliary systems.
- *Coherent attacks:* This is an attack that is in no way limited in what Even is allowed to do.

Any attack will give rise to errors in the transmission. This the *quantum bit error rate* $Q \geq 0$ captures the rate of errors in transmission. A key quantity used in security proofs is that of the *quantum mutual information* between Alice and Bob, as well as Alice and Eve and Bob and Eve. The quantum mutual information of a bipartite state defined on $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ is given by

$$I(A:B) = S(\rho_A \otimes \rho_B) - S(\rho). \tag{108}$$

This quantity captures correlations in quantum states (both classical and quantum correlations, i.e., entanglement), taking a zero value for product states.

7.2 General strategies of security proofs

Historically, the discussion of security of quantum key distribution protocols centred around the discussion of specific attacks. While this is instructive, it falls short of the actual promise of quantum key distribution. The first security proofs that considered an unbounded adversary (and hence coherent attacks) were given more than a decade after the introduction of the first schemes [7, 40, 41, 57]. Ref. [57] is noteworthy in this respect, in that it takes a very physical approach and links the theory of security of quantum key distribution to that of quantum error correction and entanglement distillation explained above.

Only again much later, it was noticed that the security criterion used so far may well be insufficient [37]: It does guarantees that an eavesdropper cannot guess the key, so in this sense the scheme is secure. But this is only true of the key is not used subsequently. If part of the key is ultimately to an eavesdropper (e.g., when it is used to encrypt a message that is known to her), the rest may become insecure. Based on these insights, a more stringent security criterion for quantum key distribution has been introduced, concomitant with new security proofs [51]. If $\rho_{K,E}$ is he joint state of the final key generated and the quantum information gathered by an eavesdropper Eve, then this state must be close to an ideal key τ_K which is perfectly uniform and is independent from the adversary's information ρ_E, as

$$(1 - p_{\text{abort}}) D(\rho_{K,E}, \tau_K \otimes \rho_E) \leq \varepsilon \tag{109}$$

where p_{abort} is the probability that the protocol aborts, $D(.,.)$ is again the trace-norm distance defined in Eq. (72) and $\varepsilon \in [0, 1]$ is a small real number. A very readable account of this development can be found in Ref. [48].

References

[1] M. Agrawal, N. Kayal, and N. Saxena. PRIMES is in P. *Ann. Math.*, pages 781–793, 2004.

[2] K. M. R. Audenaert and S. L. Braunstein. On strong superadditivity of the entanglement of formation. *Comm. Math. Phys.*, 246:443–452, 2004.

[3] K. M. R. Audenaert and B. De Moor. Optimizing completely positive maps using semidefinite programming. *Phys. Rev. A*, 65:030302, Feb 2002.

[4] K. M. R. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, and B. De Moor. Asymptotic relative entropy of entanglement. *Phys. Rev. Lett.*, 87:217902, Nov 2001.

[5] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, volume 175, page 8, New York, 1984.

[6] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, 1993.

[7] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury. A proof of the security of quantum key distribution. Proc. 32nd Symp. Th. Comp. STOC ?00, pages 715–724. ACM, 2000.

[8] E. Biham and A. Shamir. Differential cryptanalysis of des-like cryptosystems. *J. Crypt.*, 4:3–72, 1991.

[9] A. Bouchet. Connectivity of isotropic systems. *Prof. Third Int. Conf. Comb. Math.*, pages 81–93, 1989.

[10] A. Bouchet. An efficient algorithm to recognize locally equivalent graphs. *Combinatorica*, 11:315–329, 1991.

[11] S. Boyd and L. Vanderberghe. *Convex optimization.* Cambridge University Press, Cambridge, 2004.

[12] R. Cleve, D. Gottesman, and H.-K. Lo. How to share a quantum secret. *Phys. Rev. Lett.*, 83:648, 1999.

[13] M. Curty, M. Lewenstein, and N. Lütkenhaus. Entanglement as a precondition for secure quantum key distribution. *Phys. Rev. Lett.*, 92:217903, May 2004.

[14] K. K. Dabrowski, F. Dross, J. Jeong, M. M. Kanté, O.-J. Kwon, S.-I. Oum, and D. Paulusma. Recognizing small pivot-minors. In *42nd Conference on Very Important Topics (CVIT 2016)*, 2016.

[15] J. Daemen and V. Rijmen. AES Proposal: Rijndael. 2003.

[16] A. Dahlberg, J. Helsen, and S. Wehner. How to transform graph states using single-qubit operations: computational complexity and algorithms. 2018. arXiv:1805.05306.

[17] A. Dahlberg, J. Helsen, and S. Wehner. On the localizable multipartite entanglement in graph states. In *preparation*, 2018.

[18] D. Dieks. Communication by epr devices. *Phys. Lett. A*, 92:271–272, 1982.

[19] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Complete family of separability criteria. *Phys. Rev. A*, 69:022308, Feb 2004.

[20] W. Dür, J. Calsamiglia, and H.-J. Briegel. Multipartite secure state distribution. *Phys. Rev. A*, 71:042336, 2005.

[21] J. Eisert, P. Hyllus, O. Gühne, and M. Curty. *Phys. Rev. A*, 70:062317, 2004.

[22] M. Epping, H. Kampermann, and D. Bruß. Large-scale quantum networks based on graphs. *New J. Phys.*, 18:053036, 2016.

[23] M. Epping, H. Kampermann, and D. Bruß. Robust entanglement distribution via quantum network coding. *New J. Phys.*, 18:103052, 2016.

[24] C. A. Fuchs. Information gain vs. state disturbance in quantum theory. quant-ph/9611010.

[25] S. Gharibian. Strong np-hardness of the quantum separability problem. *Quantum Inf. Comp.*, 10:343–360, 2010.

[26] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *Rev. Mod. Phys.*, 74:145–195, 2002.

[27] L. Gurvits. pages 10–19, New York, 2003. ACM Press.

[28] F. Hahn, A. Pappa, and J. Eisert. Quantum network routing and local complementation. arXiv:1805.04559.

[29] M. B. Hastings. A counterexample to additivity of minimum output entropy. *Nature Phys.*, 5:255, 2009.

[30] M. Hein, W. Duer, J. Eisert, R. Raussendorf, M. Van den Nest, and H.-J. Briegel, 2006. quant-ph/0602096.

[31] M. Hein, J. Eisert, and H. J. Briegel. Multi-particle entanglement in graph states. *Phys. Rev. A*, 69:062311, 2004.

[32] M. Hillery, V. Bužek, and A. Berthiaume. Quantum secret sharing. *Phys. Rev. A*, 59:1829–1834, 1999.

[33] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Prob. Inf. Tr.*, 9:177, 1973.

[34] A. S. Holevo. The capacity of quantum channel with general signal states. *IEEE Trans. Inf. Th.*, 44, 1998.

[35] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein. Volume of the set of separable states. *Phys. Rev. A*, 58:883–892, Aug 1998.

[36] L. M. Ioannou, B. C. Travaglione, D. Cheung, and A. K. Ekert. Improved algorithm for quantum separability and entanglement detection. *Phys. Rev. A*, 70:060303, Dec 2004.

[37] Robert König, Renato Renner, Andor Bariska, and Ueli Maurer. Small accessible quantum information does not imply security. *Phys. Rev. Lett.*, 98:140502, Apr 2007.

[38] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki. Optimization of entanglement witnesses. *Phys. Rev. A*, 62:052310, Oct 2000.

[39] H.-K. Lo and H. F. Chau. Unconditional security of quantum key distribution over arbitrary long distances. *Science*, 283:2050–2056, 1999.

[40] D. Mayers. Quantum key distribution and string oblivious transfer in noisy channels. In *Adv. Crypt. CRYPTO ?96*, volume 1109 of *Lecture Notes in Computer Science*, pages 343–357, Berlin, 1996. Springer.

[41] D. Mayers. Unconditional security in quantum cryptography. *J. ACM*, 48:351–406, 2001.

[42] A. Muller, H. Zbinden, and N. Gisin. Underwater quantum coding. *Nature*, 378:449, 1995.

[43] A. Muller, H. Zbinden, and N. Gisin. Quantum cryptography over 23 km in installed under-lake telecom fibre. *Europhys. Lett.*, 33:335, 1996.

[44] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information.* Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.

[45] S.-I. Oum. Rank-width and vertex-minors. *J. Comb. Th. B*, 95:79–100, 2005.

[46] S.-I. Oum and P. Seymour. Approximating clique-width and branch-width. *J. Comb. Th. B*, 96:514–528, 2006.

[47] S. Pirandola, J. Eisert, C. Weedbrook, A. Furusawa, and S. L. Braunstein. Advances in quantum teleportation. *Nature Phot.*, 9:641–652, 2015.

[48] C. Portmann and R. Renner. Cryptographic security of quantum key distribution. arXiv:1409.3525.

[49] B. Qi, L. Qian, and H.-K. Lo. A brief introduction of quantum cryptography for engineers. arXiv:1002.1237.

[50] R. Raussendorf, D. E. Browne, and H. J. Briegel. Measurement-based quantum computation on cluster states. *Phys. Rev. A*, 68:022312, 2003.

[51] R. Renner and R. König. Universally composable privacy amplification against quantum adversaries. In *Th. Crypt. Proc. TCC 2005*, volume 3378 of *Lect. Notes Comp. Sc.*, pages 407–425. Springer, 2005.

[52] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56:131–138, Jul 1997.

[53] A. Shamir. How to share a secret. *Comm. ACM*, 22:612, 1979.

[54] C. E. Shannon. Communication theory of secrecy systems. *Bell Syst. Tech. J.*, 28:656–715, 1949.

[55] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. *Proc. 50th Ann. Symp. Found. Comp. Sc.*, 1994.

[56] P. W. Shor. Equivalence of additivity questions in quantum information theory. *Comm. Math. Phys.*, 246:473–473, 2004.

[57] P. W. Shor and J. Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85:441–444, 2000.

[58] S. Singh. *The code book.* Doubleday, New York City, 1999.

[59] M. Van den Nest, J. Dehaene, and B. De Moor. Efficient algorithm to recognize the local clifford equivalence of graph states. *Phys. Rev. A*, 70:034302, 2004.

[60] M. Van den Nest, J. Dehaene, and B. De Moor. Graphical description of the action of local clifford transformations on graph states. *Phys. Rev. A*, 69:022316, 2004.

[61] G. S. Vernam. Cipher printing telegraph systems for secret wire and radio telegraphic communications. *J. Amer. Inst. Elec. Eng.*, 45:109–115, 1926.

[62] M. Walter, D. Gross, and J. Eisert. Multi-partite entanglement. arXiv:1612.02437.

[63] J. Watrous. Simpler semidefinite programs for completely bounded norms.

[64] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, Oct 1989.

[65] S. Wiesner. Conjugate coding. *SIGACT News*, 15:78–88.

[66] M. M. Wilde. Quantum information theory. 2013.

[67] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982.